

Evaluating interviews which search for the truth with suspects:
but are investigators' self-assessments of their own skills truthful
ones?

Item type	Article
Authors	Walsh, Dave; King, Mick; Griffiths, Andy
Citation	Walsh, D. et al (2017) 'Evaluating interviews which search for the truth with suspects: but are investigators' self-assessments of their own skills truthful ones?', Psychology, Crime & Law, 23 (7):647
DOI	10.1080/1068316X.2017.1296149
Publisher	Taylor and Francis
Journal	Psychology, Crime & Law
Rights	Archived with thanks to Psychology, Crime & Law
Downloaded	19-Mar-2018 15:47:50
Link to item	http://hdl.handle.net/10545/621945

Evaluating interviews which search for the truth with suspects: but are investigators' self-assessments of their own skills truthful ones?

Dave Walsh, Mick King and Andy Griffiths

Department of Social Sciences, University of Derby, Derby, UK

ABSTRACT

Self-evaluation of one's own performance has been found in prior research to be an enabler of professional development. The task of evaluation is also a core component of a model of the investigative interviewing of victims, witnesses and suspects, being increasingly used throughout the world. However, it remains the case that there has been little research as to how practitioners approach the task itself. The present study examined the topic through the lens of observing how effectively 30 real-life investigators in the UK undertook evaluation of their interviews, representing almost the entire investigative frontline workforce of a small law enforcement agency in this country. Using an established scale of measurement, both investigators' and an expert's ratings of the same sample of interviews were compared across a range of tasks and behaviours. It was found that in almost all the assessed behaviours, requiring of the investigators to provide a self-rating, their scores tended to significantly outstrip those applied to the sample by the expert. Reasons are explored for the investigators' overstated assessments. Implications for practice are then discussed.

ARTICLE HISTORY

Received 11 August 2016
Accepted 10 February 2017

KEYWORDS

Investigative interviewing;
PEACE model; skills
evaluation; self-evaluation;
investigation skills

Introduction

Investigators from many law enforcement agencies in England and Wales are trained to interview suspects following a particular framework called the PEACE model, being an acronym for the model's component phases (that is, **P**reparation and **P**lanning, **E**ngage and **E**xplain, **A**ccount, **C**losure and **E**valuation; for a full description of the model, see Shepherd & Griffiths, 2013). Steadily, the PEACE model, based as it is on the premise of gathering reliable information rather than being focussed almost exclusively in trying to prompt confessions from suspects, has been embraced across Europe and beyond (e.g. the PRICE model in Scotland, and also adapted models in France, Ireland, the Netherlands and Scandinavia, see Walsh, Oxburgh, Redlich, & Myklebust, 2016 for more on the recent international developments of the model).

Regardless of these variations, common to all models is the **E**valuation task, comprising three functions (Walsh et al., 2016). Firstly, the investigator, during the interview, should

undertake an ongoing evaluation, determining whether the interview is meeting certain pre-planned objectives. Second, there should be a case evaluation, which assesses whether more information/evidence or enquiries are required to progress the investigation. Third, as a matter of continuing professional development, the interviewer is encouraged to reflect on his/her own performance.

Shepherd and Griffiths (2013) remark, however, that while the first two elements noted above are essential investigative tasks that should be carried out as part of every criminal investigation, the third is a specific self-development activity. Professional development evaluation should be achieved through a deliberate focus on particular elements of personal behaviour identified as requiring development based on previous personal evaluations, as an ongoing process. The idea being that individuals are to aim for incremental and continual improvement of their own performance through regular and objective reflection, further assisted by their peers and supervisors. Shepherd and Griffiths also argue that when the PEACE model was initially implemented in England and Wales during the 1990s, evaluation was regarded as too complex an undertaking. Thus, little attention was then paid to the task in terms of its utilisation as a means of skills enhancement (through, reflection and evaluation). As such, more than 20 years after the model's introduction, the present study examines whether (and if so, how well) investigators now assess their own performance accurately. As such, the present study focusses upon developmental evaluation, but not operational evaluation.

Evaluation of investigative interviews

Frequent supervisory feedback has been highlighted as important in both maintaining and improving interview performance (Lamb et al., 2000, 2002). However, in practice, interview evaluation has been found to be conducted infrequently by supervisors (Clarke & Milne, 2001; Walsh & Milne, 2007). Stockdale (1993) found, when examining the supervision of police interviews in England and Wales, that the task was undertaken irregularly, if at all (similarly, Collier & Styles-Power, 1998; Elliston, 1995; Paisley, 1998; Rigg, 1999; Stevens, 1998). Clarke and Milne's (2001) police study found an uneven approach to the supervision of interviews, with some of the 43 police forces in England and Wales either not having a supervision policy, or (where a policy existed) not attending to it. Where supervision was being applied, Clarke and Milne occasionally found supervisors either selecting an interview of brief duration for analysis, or allowing their staff to choose their own specimen interview for such supervisory evaluation.

More recently, Walsh and Bull (2011) found no single approach being applied to the task of evaluation in their study of fraud investigators in the UK. Supervisors were found inconsistent with their approach to the task with some stating that they sat in on interviews conducted by their subordinates, while others examined case files containing interview transcripts. Yet others examined tape recordings of interviews with suspects (such recordings are mandatory in England and Wales). They also stated, regardless of their chosen approach, that they conducted such evaluations at varying interval frequencies. How often their staff's performance was monitored was found largely dependent upon their perceptions of their subordinates' skill levels, with those investigators felt to be more skilful being monitored less commonly than their apparently lesser skilled colleagues.

Walsh and Bull (2011) also found that none of the supervisors in their study had an evaluation framework to structure their assessments, leading to concerns as to what competences were being assessed, and how consistently any assessment was applied. Earlier, Cherryman (2000) had found disagreement between police supervisors when they defined 'good' interviews, with some associating the definition with the gaining of a confession. We argue that the focus on such outcomes as a determinant of interview quality adversely affects the development of interviewing skills across the globe. Griffiths (2008) argues that the quality of interviews should be defined by the attempts to gain a reliable and comprehensive account, being the central purpose of the PEACE model (rather than just proving guilt or gaining a confession).

Walsh and Milne (2007) found in their survey of investigators that only a minority said they evaluated their own interview performance, also finding that they disagreed as to what was skilled interviewing. For example, some investigators tended to view the likelihood of gaining a prosecution as a measure of quality (apparently overlooking the matter that establishing a person's innocence also involves skilled interviewing). At the same time, Walsh and Milne found that some of those who did evaluate their own interviews stated that such an undertaking benefitted their interview skills (also Walsh & Bull, 2011).

Walsh and Milne (2007) also found that experienced investigators, while rarely receiving supervisory feedback on their performance, viewed themselves as skilled interviewers (regardless as to whether they were PEACE-trained or not). This finding contrasted with studies of their actual interview practice, which frequently found such investigators performing unsatisfactorily (Walsh & Bull, 2010; Walsh & Milne, 2008). Kruger and Dunning (1999) suggest that the absence of negative feedback (where due) is associated with a lack of learning and the persistence of under-performance. Other studies have also found that investigation professionals challenged when attempting to measure their own interview skills. For example, Powell, Wright and Hughes-Scholes (2011) found that police officers rated their own interviews more highly than either testimony experts or lawyers, who had assessed the same sample. However, in a more recent study, where investigators assessed their colleagues' interviews, an increased concordance of views was found between investigators, lawyers and mock jurors as to the interview quality (Ridley, Van Rheede, & Wilcock, 2015). As such, investigators may be more accurate in assessing the interview skills of their colleagues than when they are judging their own.

Self-evaluation as a task

In more general contexts, Argyris and Schön (1978) stated that evaluating one's own professional performance is a demanding task, highlighting the gap between how professionals rate themselves and how well they actually accomplish their duties. However, persistence in undertaking evaluation enabled teachers to self-assess both more accurately and objectively (Hanrahan & Isaacs, 2001; Stefani, 1998). Several studies have noted the importance of identifying and understanding skilled performance that avoids exaggerating personal strengths (Andrade & Boulay, 2003; Dunning & McElwee, 1995; Dunning, Leuenberger, & Sherman, 1995; Dunning, Meyerowitz, & Holzberg, 1989; Hayes & Dunning, 1997; Klenowski, 1995; Mamede, Schmidt, & Penaforte, 2008; Ruth-Sahd, 2003).

With regard to the domain of criminal investigation, Dando, Wilcock, and Milne (2008) found that many inexperienced volume crime police investigators felt themselves to be insufficiently trained to conduct interviews with witnesses, findings which were largely replicated in field studies of their practice (Clarke & Milne, 2001). In contrast, Bull and Cherryman (1996) found that highly experienced detectives often rated themselves as skilled, with none admitting to be 'poor'. However, Bull and Cherryman also found that many survey respondents more critical of their peers. La Rooy, Lamb, and Memon (2011) also found experienced police officers highly confident concerning their own skill levels, despite only a few having received supervisory feedback or any refresher training (see also Clarke & Milne, 2001). Such factors are felt to be crucial in maintaining skilled levels of professional performance (Lamb et al., 2000; Lamb, Sternberg, Orbach, Esplin, & Mitchell, 2002; Powell, 2002). Griffiths (2008) has also found that training alone is insufficient to sustain skill levels. He found that training programmes lacking subsequent reinforcement activities (such as supervisory monitoring and/or refresher training initiatives) are associated with significant declines in performance over time.

In field studies that concern assessment of actual performance of interviewer skills (e.g. Clarke & Milne, 2001; Walsh & Bull, 2010; 2015; Walsh & Milne, 2008) there were few skilled examples, and it was often the case that unacceptable levels of performance were found. It is possible therefore that the noted erroneous self-assessment of skills by investigators might well be related to their own lack of ability. Dunning, Johnson, Ehrlinger, and Kruger (2003) have argued that incompetence is inevitably associated with inflated self-assessment since it is improbable to expect individuals, yet to be competent, to understand how such competence is represented in practice.

Somewhat counter-intuitively however, (at least on first blush), Dunning et al. also found that highly skilled performers inaccurate in their own self-assessments. However, they found that such assessments were not exaggerated, but were ones made relative to those who were inferior performers. That is, the higher skilled depressed their self-assessments through over-estimated perceptions of actually less effective peers (whom they had over-assessed). Greater accuracy in their own skills was only achieved when these high performers were exposed to the actual skill levels of those less capable (whom they then recognised as such). However, such exposition to competent performance did not always have such positive effects on accuracy on those lacking competence (Kruger & Dunning, 1999). When more accurate self-assessment occurred, it was found likely associated with those who responded positively to training that improved their metacognitive skills and logical reasoning. For others, their incompetence (and also their over-estimation of their own ability) stubbornly remained.

Several authors (e.g. Alicke & Sedikides, 2009; Baccus, Baldwin, & Packer, 2004; Dijksterhuis, 2004; Taylor, Lerner, Sherman, Sage, & McDowell, 2003) have reported the benefits to individuals of positive self-perception in aiding their own psychological well-being. Various phenomena and motivational dispositions lead people to assess whether the impressions they convey are likely to lead to social acceptance or rejection, importantly protecting their own self-esteem. Unsurprisingly, therefore, several previous studies (e.g. Alicke & Govorun, 2005; Sedikides & Gregg, 2003, 2008) have found that people tend to overstate their own merits, while overlooking their shortcomings. Indeed, Taylor (2014) provides further evidence of self-enhancement (even in the face of most obvious contrary evidence). In her study, participants continued to rate themselves as both 'skilful' burglars

(when assessing how they carried out their criminal activity) and decision-makers (when judging when and where to commit their crimes), despite all her study participants having been convicted of these offences!

A self-serving bias has been argued to be a motivation for such incorrect evaluations and misattributions, with people's recall invariably re-constructed to place their own attributes in a manner more advantageous to their own self-interests (Campbell & Sedikides, 1999; Mezulis, Abramson, Hyde, & Hankin, 2004; Roese & Olson, 2007; Sedikides, Campbell, Reeder, & Elliot, 2002). For example, successful outcomes are attributed to dispositional factors (such as their own ability and effort), arguably taking more credit than is deserved (Alicke & Sedikides, 2009). Unsuccessful outcomes are, in turn, ascribed to situational factors (such as bad luck, task complexity or being due to 'others') (Weiner, 1972). In addition to such displacement strategies, other reasoning may include (i) denying shortcomings; (ii) employing compensatory qualities to both offset failure and rationalise the situation and (iii) comparing themselves to 'lesser' others to gain a more favourable self-perception (Sedikides & Gregg, 2003). As such, these deflections assist self-protection, particularly where any negative feedback threatens either their own interests or peer approval (Alicke & Sedikides, 2009). Meanwhile, Sedikides, Green, and Pinter (2004) found their participants only selectively recalling feedback (particularly when that feedback appeared to threaten their self-perceptions). These authors found this phenomenon taking place even when such threatening feedback was consistent with that which their participants had earlier described as being unfavourable. As Sedikides and Gregg (2006) pronounce, self-protection motives considerably outstrip those pertaining to ones of self-assessment and self-verification.

Weinstein (1984) found that people selectively recall their own strengths (and others' weaknesses), giving themselves more credit as perceived positive actions, attributes and efforts are more easily brought to mind than negative ones (Alicke & Sedikides, 2009; Skowronski, Betz, Thompson, & Shannon, 1991; Walker, Skowronski, & Thompson, 2003). One reason might be that individuals, when reflecting, find it more demanding to prompt negative self-inferences than negative other-inferences since they invariably possess more positive knowledge about themselves than about others (Kihlström, Beer, & Klein, 2003; Prentice, 1990). Moreover, these positive recollections decay at a slower rate than their less favourable counterparts (Skowronski, Gibbons, Vogl, & Walker, 2004; Walker et al., 2003). As such, positive information is elevated (or reconstructed as positive) in one's hierarchy of memory, and self-threatening feedback is derogated (Baumeister, Campbell, Krueger, & Vohs, 2003; Sedikides & Gregg, 2008; Sedikides, Gaertner, & Toguchi, 2003), leaving heuristically driven information to result in biased judgments (Kahneman, Slovic, & Tversky, 1982).

Sedikides, Green and Pinter (2004, also Sedikides & Green, 2006) refer to the confirmation bias process, where individuals focus on information that self-affirms or self-enhances (and fail to consider that which does not) in efforts to self-protect, as the 'mnemonic neglect model'. As such, any new information, inconsistent with the existing and biased information, has to be set against a larger and more established body of knowledge, making it harder for this newer and smaller amount of information to alter existing, and longer held judgements (even when it is known their own performance is being assessed by another, see Preuss & Alicke, 2009). However, when assessing third parties, since one will likely possess less existing awareness of them, the same challenges do

not tend to arise. (Van Overwalle & Labiouse, 2004). As such, humans tend to, by and large, evaluate themselves either more favourably than the objective facts evidence (Gosling, John, Craik, & Robins, 1998), or than third party observers think justified (Epley & Dunning, 2000). Moreover, people have been found to rate themselves more favourably than they rate their peers on the basis of identical behavioural evidence, including that evidence which they pronounce as negative when it is manifest in others (Alicke, Vredenburg, Hiatt, & Govorun, 2001).

Such misperceptions can lead to opportunities for self-improvement to be missed (Dunning, Heath, & Suls, 2004; Duval & Silvia, 2002; Oettingen & Gollwitzer, 2001; Wilson & Dunn, 2004). Accurate reflections of their own performance are necessary when professionals evaluate their own skill (Boud, Cressey, & Docherty, 2006; Brown, Harris, & Harnett, 2012; Cropley, Miles, Hanton, & Niven, 2007; Mamede et al., 2008; Moon, 2006; Sheikh, Milne, & MacGregor, 2007). Self-protectionist attitudes can lead to various positive states (e.g. subjective well-being, optimism, active coping), while avoiding negative ones (e.g. depression, anxiety, neuroticism) (Dunning, 2005; Marshall & Brown, 2007; Sedikides, Gregg, & Hart, 2008). However, if self-protection becomes entrenched, so may failure to learn and develop from one's range of experiences (Colvin & Griffo, 2007; Sedikides, 1999; Sedikides & Luke, 2007).

It has, however, been found that skills development is likely to occur when feedback is received from those perceived as trusted (or credible) observers (Sedikides, Campbell, Reeder, & Elliot, 2002). Such a situation demands that the individual acknowledges a momentarily deflated self-image in order to gain self-enhancement in the longer term (Sedikides & Strube, 1997). Prior studies (e.g. Buunk, Cohen-Schotanus, & Henk van Nek, 2007; Lockwood & Kunda, 1997) have found that short-term threats may be accepted, if there is perceived to be a reasonable prospect of long-term improvement by learning from initially threatening feedback. In turn, this may overcome the selective processing and remembering effects of the 'mnemonic neglect model'.

Opportunities for self-improvement can also occur through self-evaluation when using comparative, rather than absolute, judgements (Sedikides & Gregg, 2003). Social juxtapositioning of skills becomes self-protectionist (Crocker, Voelkl, Testa, & Major, 1991), when solely contrasted to perceived inferior counterparts that maximises an *ego-defensive contrast effect* (Biernat & Billings, 2001; Suls & Wills, 1991). However, comparisons, when conducted in an upward direction, have been found to instigate positive developments (Gruder, 1971; Miller, Turnbull, & McFarland, 1988). This has been found to be particularly the case where such relationships are viewed as (i) strong (Collins, 1996); (ii) uncompetitive (Wood, 1989) and (iii) not so significantly inferior that any gap cannot likely be bridged (Lockwood & Kunda, 1997; Wheeler, 1996).

The foregoing portrays how inaccurate self-evaluations may both emerge and endure. Nevertheless, enabling professionals to self-evaluate can prompt their development (Goodyear, 1997; Moon, 2006; Schon, 1987). As such, the present study examines how accurately investigators evaluate their own interviewing performance, based on dimensions of performance integral to the PEACE model. We hypothesised, based on human tendencies to self-enhance (extensively noted in the foregoing literature review), that investigators would rate themselves more highly on these dimensions when compared with an independent expert's assessments.

Methodology

Initial procedure

Investigators working for a law enforcement agency in the UK, whose then total frontline investigation strength of 36, all attended a three-day refresher training course between November 2013 and March 2014. The course was delivered (and accredited) by the authors' home University. As part of the accreditation process, trainees each had to undertake at least one mock interview with stooges acting as suspects of offences within the agency's jurisdiction of labour exploitation, which the investigators would later self-assess. Although undertaking the interviews was an expectation of their duty, the investigators were given the choice in writing whether their interviews were to be allowed to go forward to be used as research data for the present study. All investigators (except one) complied with the researchers' request, providing their informed consent, leaving 35 investigators in the sample.

The 12 stooges, variously comprised university colleagues, retired and serving police officers, and actors from a local drama company, were all unknown, prior to their training, to the investigators. The training was led by the second author ('the expert'), a former police trainer with 30 years' professional experience alongside his possessing a Masters' degree in Criminal Investigation. He was accompanied from time to time during this training by the first author, who, in addition to possessing over 20 years' investigative experience, has a PhD alongside authorship of many peer-reviewed published studies in this subject area.

As in real-life, investigators conducted the mock interviews in pairs, with one nominated as the lead interviewer while the other played a much lesser role. Those interviews, where the interviewer undertook the lead role, were those used in the present study for assessment. Seven different case studies were used (spread equally across the sample). The number of cases studies ensured that none of the investigators became overly familiar with any particular case, either through undertaking other interviews (as the second interviewer) during their training or by discussion with their colleagues who had already undertaken the training. However, what was common to all seven cases was that they were typical of the investigations the trainees would have faced in their day-to-day duty. Each case study was constructed by both the researchers and the agency's National Development Manager, being given to participants (i.e. both investigators and the mock suspects) around 24 h before the interviews were due to be conducted. The details of the case study given to the stooges differed from that given to investigators, in that the stooges were told of their actual role in the case under 'investigation' (something which the investigators would aim to establish through their questioning). As such, the stooges were situated in one of three conditions. That is, they had either committed an offence or they were innocent, while others, through their lack of knowledge of the law, were unaware they had committed any illicit activity. As such, they saw themselves (at least initially) as innocent.

The mock suspects were advised by the researchers that their approach in the interview was to be co-operative. Those suspects, whose scripts meant that they had committed offences, were to adopt a denial strategy (at least initially), engaging in deception to avoid their guilt being revealed. All mock suspects were advised by the researchers that

any co-operation would only extend to the level of rapport that they felt existed between them and their interviewers.

The interviews, which were conducted on the agency's official premises, were video-recorded. The recordings would last a maximum of 45 min, although some interviews would be expected to (and did) finish sooner. Investigators were each given a copy of their recorded interview. Investigators were also given both verbal and written guidance on how to complete a self-rating scale, being asked to later assess their own interview performance of that recorded interview. As five of the interviews lasted less than 15 min, they were removed from the sample as there was insufficient data for assessment using the scale (following Walsh & Bull, 2010), leaving a sample of 30 for assessment. Of this sample, all investigators had earlier been trained in the PEACE model. In terms of their experience, 21 had conducted investigations for over 10 years, with a further 7 possessing between 5 and 10 years of experience, while of the remainder each had been serving as investigators for less than 2 years (and had each undertaken no more than 50 interviews). Of those with more than 2 years' experience, 26 stated that they had conducted more than 300 interviews, while 2 stated that they had conducted around 150–200 interviews. Twenty of the sample involved male investigators as lead interviewers.

The instrument of measurement

The rating scale had evolved from prior studies (e.g. Bull & Cherryman, 1996; Clarke & Milne, 2001; Griffiths, 2008; Walsh & Bull, 2010; Walsh & Milne, 2008, thus incorporating agreed interviewing skills, recommended in the PEACE framework). In sum, the scale possessed 38 dimensions of interview performance. The scale examined interviewer performance dichotomously (absent/present) over eight items. These were (i) identifying all present in the interview; (ii) providing time of commencement, location and date of interview; (iii) advising of right to both legal advice and copy of interview recording; (iv) and giving reason (i.e. which suspected offence) and purpose (i.e. to give their account) for interview. The scale also included an ascending 5-point Likert scale (for 30 items) that measured interview skills, where a score of '1' represented poor performance, '2' inadequate and '3' being rated as satisfactory. A score of '4' would be applied to skilled performance levels, while a rating of '5' would be provided to highly skilled examples. Following other studies (e.g. Clarke & Milne, 2001; Walsh & Bull, 2010, 2012; Walsh & Milne, 2008), a score of '3' on the scale would be the minimum viewed as acceptable.

Procedure following investigator measurement

Once the investigators had independently undertaken the self-assessment task, they dispatched their assessments within 2 weeks of the completion of the course by email to the first author, being highly familiar with the rating scale and its definitions. The expert, blind of the investigators' self-ratings, conducted his own assessments of the entire sample from a second copy of the video recordings. Before examining the self-ratings, 38 per cent ($n = 10$) of the expert's ratings were subject to measures of reliability by the first author, generally finding a high level of concordance (see Table 1). Total agreement between both experts was also found on all eight dichotomous measures of the scale. Both the

Table 1. Inter-rater reliability correlations between the two experts.

Task		Task	
Preparation and planning	0.86	Keeps interview to relevant topics	0.83
Caution quality	0.95	Supplies summaries/links	0.92
Checking understanding of caution	0.95	Covers points to prove	0.95
Informing of ongoing legal advice	1.00	Explores information received	0.80
Advising of right to tape copy	1.00	Explore any inconsistencies	0.80
Explains purpose of interview	0.98	Challenging skills	0.91
Advises routines and route map	0.95	Explores motive	0.81
Explaining interview is opportunity to give account	0.92	Uses pauses and silences	0.88
Rapport building	0.80*	Conversation management skills	0.80*
Encourages suspect to give account	0.83	Cognitive interview skills	0.86
Develops investigative topics	0.80*	Expresses self-confidence	0.91
Dealing with difficulty	0.83	Expresses open mindedness	0.91
Supplies logical interview structure	0.77*	Shows flexibility	0.77*
Appropriate questioning techniques	0.92	Communication skills	0.92
Appropriate questioning strategies	0.92	Active listening skills	0.77*

Note: All correlations significant; $p \leq .01$, except * where $p \leq .05$.

expert's and the investigators' sets of assessments were then compared to each other by the first author.

Results

The mean duration of the sample of 30 interviews was 37.03 min ($SD = 7.62$) in the range of 24–45 min. Omissions were found by the expert in the sample concerning the investigators' requirement to provide the interview (i) date ($n = 2$); (ii) time ($n = 1$); (iii) location ($n = 3$), while the right to legal advice was not explained in three interviews. In four interviews, identification of those present was not undertaken, while in six cases there was no explanation provided as to how the suspect could gain access to a copy of the tape recording of the interview. In 15 of the 19 cases in the sample (where the interview was completed within 45 minutes), the interviewer failed to ask the suspect whether he/she wished to alter or add anything to which had been said in the interview. Neither on three occasions in such interviews was the cessation time given. In five of these eight present/absent areas of assessment (i.e. giving interview date, start time, and location, advice either regarding legal representation or availability of a copy of the tape), there was complete agreement between the assessment of the expert and that of the investigators concerning the incidence rate of each task. With regard to other tasks, however, there was discord as to their frequency of presence across the sample (that is, identification of *those present*, where expert $n = 26$, investigators $n = 28$; invitation to add/alter anything; expert $n = 4$, investigators $n = 7$; stating interview cessation time; expert $n = 10$, investigator $n = 13$).

The caution (that provides, in England and Wales, the suspect with advice as to their right to silence and the evidential nature of any answers that are provided, while making a suspect aware of the adverse inference any court might take if they remain silent) was provided in all the interviews. Two breaches of the law relating to the interviewing of suspects in England and Wales was found to occur in the same interview. These involved the investigator suggesting the presence of evidence that was known not to exist, while an oppressive and accusatorial interview style was found to have been conducted in that interview.

Table 2 shows that of the 30 behaviours measured for skill levels the investigators without exception rated themselves as more skilled than was independently assessed by the expert. With the exception of just four of these assessed behaviours (that is, *preparation and planning*, *covering legal points to prove*, *expressing self-confidence*, and *open mindedness*), the difference between the two ratings was found to be statistically significant. Effect size calculations consistently showed (again with only a few exceptions) that the strength of the differences between self and expert judgements to be substantially robust. Since (according to Clarke & Milne, 2001; Walsh & Bull, 2010; Walsh & Milne, 2008) ratings of '3' and above relate to the minimum acceptable standard, Table 2 also shows that on average the investigators, with the exception of 5 of the 30 assessed behaviours, regarded themselves as attaining at least this particular standard. However, even in those five behaviours, the expert provided a mean rating that was still significantly less than investigators had so self-assessed.

In a further five categories, investigators' mean score reflected that they saw themselves as 'skilled' ($M \geq 4.00$). For example, 18 investigators gave themselves a 'highly skilled' rating of '5' in the behavioural dimension of *delivering the caution*, compared to the expert who assessed only four similarly so high in this behavioural dimension. With regard to the skills concerning *checking suspects' understanding of the caution*, 10

Table 2. T-test results across sample between expert and self-ratings (mean scores, with SD in brackets).

Behavioural dimension	Self-rating	Expert rating	<i>t</i> value	<i>d</i> value
Preparation and planning	3.57 (0.73)	3.33 (0.61)	1.35	0.35
Caution quality**	4.53 (0.63)	3.83 (0.83)	3.67	0.96
Checking understanding of caution**	4.13 (0.97)	2.80 (1.24)	4.64	1.22
Informing of ongoing legal advice**	4.03 (1.27)	2.77 (1.17)	4.02	1.06
Advising of right to tape copy**	3.73 (1.50)	1.93 (0.93)	5.51	1.45
Explains purpose of interview**	3.87 (1.11)	2.63 (1.07)	4.40	1.16
Advises routines and route map**	3.30 (1.12)	1.30 (0.60)	8.64	2.27
Explains opportunity to account**	3.37 (1.45)	1.57 (0.97)	5.65	1.48
Rapport building*	3.77 (0.77)	3.20 (0.96)	2.52	0.66
Encourages suspect to account**	3.93 (0.64)	2.93 (1.23)	3.96	1.04
Develops investigative topics**	3.77 (0.73)	3.03 (0.93)	3.41	0.90
Dealing with difficulty**	3.85 (0.67)	2.63 (0.72)	5.28	1.39
Supplies logical interview structure**	4.00 (0.79)	3.17 (0.83)	3.98	1.05
Appropriate questioning techniques*	3.80 (0.76)	3.24 (0.87)	2.62	0.69
Appropriate questioning strategies**	3.63 (0.72)	2.93 (0.87)	3.40	0.89
Keeps interview to relevant topics**	4.17 (0.75)	3.43 (0.50)	4.46	1.17
Supplies periodic summaries/links**	3.27(1.01)	1.87 (0.87)	5.76	1.51
Covers points to prove	3.23 (1.07)	2.83 (0.70)	1.71	0.45
Explores information received**	3.50 (0.82)	2.87 (0.94)	2.79	0.73
Explore any inconsistencies*	3.17 (1.12)	2.57 (0.57)	2.62	0.69
Challenging skills* #	2.77 (1.30)	1.67 (0.70)	2.41	0.75
Explores motive** # #	2.47 (1.22)	1.43 (0.69)	3.94	1.03
Uses pauses and silences**	2.77 (1.30)	1.28 (0.65)	5.52	1.45
Conversation management skills*	3.50 (1.01)	2.93 (0.64)	2.60	0.68
Cognitive interview skills**	2.97 (1.00)	1.21 (0.49)	8.53	2.24
Expresses self-confidence	3.80 (0.96)	3.60 (0.68)	0.93	0.24
Expresses open mindedness	3.80 (0.85)	3.57 (0.73)	1.14	0.30
Shows flexibility**	3.80 (0.81)	3.00 (0.57)	4.40	1.16
Communication skills**	3.97 (0.77)	3.37 (0.49)	3.62	0.95
Active listening skills**	3.83 (0.64)	3.21 (0.73)	3.50	0.92

Note: All correlations significant; $df = 58$, except # where $df = 37$, and ## where $df = 42$ (due to some interviews in the sample not demonstrating those particular behaviours).

* $p \leq .05$, ** $p \leq .01$.

investigators assessed themselves as achieving a highly skilled rating of '5', whereas the expert assessed alike on three occasions. More, while 15 investigators felt that they *explained to the suspect the right to legal advice* in a highly skilled manner, the expert only agreed on two occasions that this was the case. Further, 10 investigators felt they were highly skilled with concern to *keeping the interview to relevant topics*, in contrast to the expert (who did not rate any interviewer at this performance level). Finally, 10 investigators awarded themselves a highly skilled rating when *undertaking an appropriate interview structure*. The expert largely contradicted such high self-assessments, finding that only two interviews were so highly skilled.

As Table 2 also depicts, in 12 of the 30 rated behavioural dimensions, the expert gave a mean score of '3' (being viewed in the literature as the threshold of acceptable performance). In contrast, investigators' mean scores in 5 of those 12 categories were in the '4' range (i.e. a skilled performance). Further, Table 2 also shows that the expert assessed a mean score of '1' (i.e. poor) in 9 behavioural dimensions, whereas the investigators never rated themselves on average in any of the 30 categories at such low skill levels.

The total number of the self-assessments undertaken throughout the sample by the 30 investigators amounted to 864 individual ratings across the range of 30 behavioural dimensions (not all investigators provided a self-rating for every dimension). Of these 864 ratings, on only 46 occasions (i.e. 5.3%) were the investigators' scores lower than that of the expert, while 615 (i.e. 71.2%) of the self-ratings were higher than that of the expert. Further, 395 (64.2%) of these 615 ratings were found to be at least 2 integers higher than the corresponding expert rating (e.g. investigator self-rated a '4' score, whereas the expert assessed the same performance as '2'). Self-ratings in the highest two categories (i.e. a score of '4' or '5', being skilled or highly skilled) were found to have occurred 528 times (61.1%). In comparison, the expert gave such correspondingly high scores on just 198 occasions (22.9%). The expert also assessed the performance of investigators on 406 occasions (47%) in either of the lowest two category scores of '1' and '2' (i.e. 'poor' and 'unsatisfactory' ratings, respectively). Investigators meanwhile, scored themselves at '1' or '2' in only 115 of the 864 ratings (13.3%).

Discussion

The present study is believed to be the first, as far we know, to explicitly examine the capabilities of investigators to accurately evaluate their own interview performance, a task that is a core component of the original PEACE model. Prior field studies (e.g. Clarke & Milne, 2001; Walsh & Bull, 2012, 2015) have hitherto tended to concern examining tasks and behaviours demonstrated within interviews (such as building/maintaining rapport, undertaking questioning strategies, disclosing evidence, and avoiding confirmation bias or false confessions). However, self-evaluation has been found by educators and clinicians to be a key factor in their enhancement of the skills required in their professions.

The present study found that the main hypothesis of this study (i.e. that investigators would self-enhance) was supported. Investigators across the vast majority of those measured behaviours rated themselves significantly higher than did the expert. It was found that in only a small minority of instances did the expert provide a higher rating than that of the investigator. The majority of the mean ratings of the behavioural dimensions, assessed by the expert, fell below the level of satisfactory performance. It is therefore

particularly resonant that Dunning et al. (2003) argued that incompetent performers are unlikely to accurately self-assess (since they cannot be expected to judge competence with any precision, particularly when a definition of competence is equivocal). Such a pronouncement appears to be evident from the findings of this study. Further, investigator self-ratings, at a level defined as most skilled, were found to be more than twice the prevalence of the expert's similar ratings. In addition, investigators assessed themselves at the 'lower end' ratings of '1' and '2' on the scale less than three times the incidence rate than did the expert. Such lesser assessments have been deemed in prior studies as being beneath the threshold of acceptable performance levels (see Clarke & Milne, 2001; Walsh & Bull, 2010). The absence of timely, accurate and (where warranted) negative feedback found in previous studies of investigators (e.g. Walsh & Milne, 2007) is also thought to be contributory to the failure to recognise their own true skills (Dunning et al., 2003).

This persistent over-assessment by investigators of their own performance, found in the present study, might be explained when the behaviour of *self-confidence* is examined. This particular behaviour was one of the few where no significant difference was found between the investigators' and the expert's ratings. Not only was there statistical parity, but also the expert rated this behaviour second highest (after the procedurally repetitive task of delivering the caution). It is speculated whether such high levels of self-belief were indeed evident when it came to the task of investigators assessing their own interview performance.

Self-protection, found in other studies as a cause of self-enhancement (Alicke & Sedikides, 2009; Baccus et al., 2004; Dijksterhuis, 2004; Sedikides et al., 2004; Taylor et al., 2003), might be an explanation for the present study's findings, possibly accentuated by the scale used in the present study. Though used in prior research (e.g. Clarke & Milne, 2001; Shawyer, 2009; Walsh & Bull, 2010; Walsh & Milne, 2008), the scale might have in itself encouraged self-ratings to be inflated. That is, the rating scale through its reliance on quantitative measures, despite the investigators being trained in its use, might be better restricted to third-party assessments than self-evaluations. For example, being aware of the scale's definitions (e.g. where '1' = 'poor'), investigators may exaggerate their own scores to avoid viewing themselves as repeatedly incompetent over the 30 dimensions of assessment.

Investigative interviewing has been regularly found, in self-report studies, to be one of the most important, and one of the most regular tasks that investigators undertake in their professional roles (e.g. Bull & Cherryman, 1996; Walsh & Milne, 2007). As such, self-ratings, no matter how accurate, applied towards the lower end of the scales on a frequent basis, may be more of an admission of their failings than which investigators are prepared to accommodate or acknowledge, so central is the task to their investigative professionalism. Walsh and Milne (2007), for example, found that experienced investigators, though yet to undergo training, all rated themselves as good interviewers, similar to findings in other studies involving investigators' (often found misplaced) perceptions of their own ability (Bull & Cherryman, 1996; La Rooy et al., 2011; Powell et al., 2011; Walsh & Bull, 2011).

Heuristically motivated memorial performance, found in other self-evaluation studies (e.g. Kahneman et al., 1982; Sedikides et al., 2003; Skowronski et al., 1991, 2004; Walker et al., 2003), might be another reason for investigators' inaccurately rating themselves, allowing them to obviate exercising the required cognitive effort when self-rating. To

lessen such concerns, the researchers advised the investigators to examine the video-recording of the interview which they were to self-assess at least twice before entering their scores. We do not know whether this advice was followed. However, some reassurance is provided that they did examine the recordings since there was general agreement between them and the expert when considering the majority of the scale's eight tasks that were assessed on a dichotomous (present/absent) basis. Regardless as to whether these responses were by chance alone, index scores measured along such a 5-point scale might have encouraged some investigators to provide scores (with insufficient thought or concern as to their precision or qualification), which were not necessarily a true reflection of their displayed abilities.

Future directions

Measures, as those in the present study appear, do not appear to enable investigators to accurately capture their own performance levels. As such, we do not recommend their use in professional settings. However, we wonder whether other methods might better prompt investigators to examine their own performance in a more reflective (and truthful) manner. For example, in other settings, the usage of reflective journals has been found to develop reflective skills (Connor-Greene, 2000; Patton, Wood, & Agarenzo, 1997; Salmon, 2006; Woodward, 1998). Consistently found in these studies, in the domain of education, was that such reflective logs not only enabled evidence of understanding to be provided of their own technical skills/knowledge and of critical self-reflection/self-awareness, it was also argued to lead to performance improvements (Biggs, 1999; Connor-Greene, 2000; O'Rourke, 1998). This approach might therefore avoid any concerns associated with pre-conceived classifications (of 'poor', 'satisfactory' and 'excellent' skills), with the possibility of more truthful insights emerging from an approach that requires more cognitively strenuous effort. Indeed, Kruger and Dunning (1999) noted less capable performers improving their self-awareness of their shortfalls in skills when in tandem with increased logical reasoning and metacognition.

We also recognise that the demand characteristics from the use of mock suspects in the study might have had a confounding effect on the interview dynamics, if not the skill measures. As such, we recommend future research examines actual interviews from the field. Further (in view of the matter that the present study involved comparison of a single interview for each investigator), future samples should contain those pertaining to several interviews for each investigator to overcome the possible idiosyncrasies that might be apparent in such a narrow selection.

While prior studies have found judgments of skills influenced by apparent interview outcome, we made no attempt to account for this as 11 of the interviews were curtailed on time grounds and therefore did not conclude. As such, including the outcome as a variable would have reduced the sample yet further. Further, the skills measured in the present study are those that constitute the bedrock of the PEACE model, where the aim is singularly to obtain fulsome and reliable accounts (rather than, say, untested confessions). As such, we were only concerned with measurement of those skills that are consistent with the aims of the PEACE model and no others. Nevertheless, we recognise that prior studies have found that interviews which are oriented towards outcomes as confessions or proving guilt, tainted investigators' views of skill levels. To determine whether such beliefs

persist, we recommend that future research incorporates such measures to identify their having any influence upon inflated self-assessments.

The present study suggests that investigative interviewing, while developing a more ethical means of questioning suspects, still continues to be blighted by mediocre performance levels, found often in studies of actual practice (Clarke & Milne, 2001; Griffiths & Milne, 2006; Walsh & Bull, 2010; Walsh & Milne, 2008).

Concluding thoughts

The present study has found that investigators tend to be self-enhancing when being asked to evaluate their own performance, much in line with the volume of literature concerning self-assessment (whether or not affiliated to the domain of criminal investigation). It is also argued that, despite near a quarter of a century after the introduction of the PEACE model, many investigators do not appear to customarily undertake evaluations of their interviews. They are not helped when such ambiguity exists (at least in practitioner communities) as to what constitutes good interviewing skills. Yet this awareness does exist (Clarke & Milne, 2001; Griffiths & Milne, 2006; Walsh & Bull, 2010, 2015). Upon being asked to actually perform this task, this present study's findings suggest they seldom do evaluate with either reliability or validity. It therefore remains doubtful whether they undertake a systematic approach, or even (with reference to the scientific literature) one that follows both agreed and understood standards. This is despite the task of evaluation, as a means of professional development, being one of the model's fundamental elements. Investigators have been found to be highly favourable towards the current investigative interviewing framework of the PEACE model (Walsh & Milne, 2007). However, we contemplate, in light of the present study's findings, whether such judgments by professionals is a less a case of endorsing PEACE and more one of undertaking PEAC.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

- Alicke, M. D., & Govorun, O. (2005). The better-than-average effect. In M. D. Alicke, D. A. Dunning, & J. I. Krueger (Eds.), *The self in social judgment* (pp. 85–106). Philadelphia, PA: Psychology Press.
- Alicke, M. D., & Sedikides, C. (2009). Self-enhancement and self-protection: What they are and what they do. *European Review of Social Psychology*, 20, 1–48. Retrieved from <http://dx.doi.org/10.1080/10463280802613866>
- Alicke, M. D., Vredenburg, D. S., Hiatt, M., & Govorun, O. (2001). The “better than myself effect”. *Motivation and Emotion*, 25, 7–22.
- Andrade, H., & Boulay, B. (2003). Role of rubric-referenced self-assessment in learning to write. *Journal of Educational Research*, 97(1), 21–34. Retrieved from <http://dx.doi.org/10.1080/00220670309596625>
- Argyris, C., & Schön, D. (1978). *Organizational learning: A theory in action perspective*. New York, NY: Addison-Wesley.
- Baccus, J. R., Baldwin, M. W., & Packer, D. J. (2004). Increasing implicit self-esteem through classical conditioning. *Psychological Science*, 15, 498–502. Retrieved from <http://dx.doi.org/10.1111/j.0956-7976.2004.00708.x>

- Baumeister, R. F., Campbell, J. D., Krueger, J. I., & Vohs, K. D. (2003). Does high self-esteem cause better performance, interpersonal success, happiness, or healthier lifestyles? *Psychological Science in the Public Interest*, 4, 1–44. Retrieved from <http://dx.doi.org/10.1111/1529-1006.01431>
- Biernat, M., & Billings, L. S. (2001). Standards, expectancies, and social comparison. In A. Tesser & N. Schwartz (Eds.), *Blackwell handbook of social psychology: Intraindividual processes* (pp. 257–283). Oxford: Blackwell. Retrieved from <http://dx.doi.org/10.1002/9780470998519.ch12>
- Biggs, J. (1999). *Teaching for quality learning at university*. Buckingham: Society for Research into Higher Education & Open University Press.
- Boud, D., Cressey, P., & Docherty, P. (2006). Setting the scene for productive reflection. In D. Boud, P. Cressey, & P. Docherty (Eds.), *Productive reflection at work* (pp. 3–10). London: Routledge.
- Brown, G. T. L., Harris, L. R., & Harnett, J. (2012). Teacher beliefs about feedback within an assessment for learning environment: Endorsement of improved learning over student well-being. *Teaching and Teacher Education*, 28, 968–978. Retrieved from <http://dx.doi.org/10.1016/j.tate.2012.05.003>
- Bull, R., & Cherryman, J. (1996). *Helping to identify skills gaps in specialist investigative interviewing: enhancement of professional skills*. London: Home Office.
- Buunk, A. P., Cohen-Schotanus, J., & Henk van Nek, R. (2007). Why and how people engage in social comparison while learning social skills in groups. *Group Dynamics: Theory, Research, and Practice*, 11, 140–152. Retrieved from <http://dx.doi.org/10.1037/1089-2699.11.3.140>
- Campbell, K. W., & Sedikides, C. (1999). Self-threat magnifies the self-serving bias: A meta-analytic integration. *Review of General Psychology*, 3, 23–43. Retrieved from <http://dx.doi.org/10.1037/1089-2680.3.1.23>
- Cherryman, J. (2000). *Police investigative interviewing: Skill analysis and concordance of evaluations* (Unpublished PhD thesis). University of Portsmouth, Portsmouth.
- Clarke, C., & Milne, R. (2001). *A national evaluation of the PEACE Investigative Interviewing Course*. London: Home Office.
- Collier, F., & Styles-Power, C. (1998). *An evaluation to determine the extent to which investigative interviewing techniques and ethics have been incorporated into working practices within Cambridgeshire Constabulary*. Suffolk: The Collier Consultancy.
- Collins, R. L. (1996). For Better for worse: The Impact of upward social comparison on self-evaluation. *Psychological Bulletin*, 119, 51–69. Retrieved from <http://dx.doi.org/10.1037/0033-2909.119.1.51>
- Colvin, C. R., & Griffo, R. (2007). On the psychological costs of self-enhancement. In E. C. Chang (Ed.), *Self-criticism and self-enhancement: Theory, research, and clinical implications* (pp. 19–35). Washington, DC: American Psychological Association. Retrieved from <http://dx.doi.org/10.1037/11624-008>
- Connor-Greene, P. A. (2000). Assessing and promoting student learning: Blurring the line between teaching and testing. *Teaching of Psychology*, 27, 84–88. Retrieved from http://dx.doi.org/10.1207/s15328023top2702_01
- Crocker, J., Voelkl, K., Testa, M., & Major, B. (1991). Social stigma: The affective consequences of attributional ambiguity. *Journal of Personality and Social Psychology*, 60, 218–28. Retrieved from <http://dx.doi.org/10.1037/0022-3514.60.2.218>
- Cropley, B., Miles, A., Hanton, S., & Niven, A. (2007). Improving the delivery of applied sport psychology support through reflective practice. *The Sport Psychologist*, 21, 475–494.
- Dando, C., Wilcock, R., & Milne, R. (2008). Victims and witnesses of crime: Police officers' perceptions of interviewing practices. *Legal & Criminological Psychology*, 13, 59–70. Retrieved from <http://dx.doi.org/10.1348/135532506X162498>
- Dijksterhuis, A. (2004). I like myself but I don't know why: Enhancing implicit self-esteem by subliminal evaluative conditioning. *Journal of Personality and Social Psychology*, 86, 345–355. Retrieved from <http://dx.doi.org/10.1037/0022-3514.86.2.345>
- Dunning, D. (2005). *Self-insight: Roadblocks and detours on the path to knowing thyself*. New York, NY: Psychology Press. Retrieved from <http://dx.doi.org/10.4324/9780203337998>
- Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: Implications for health, education, and the workplace. *Psychological Science in the Public Interest*, 5, 69–106. Retrieved from <http://dx.doi.org/10.1111/j.1529-1006.2004.00018.x>

- Dunning, D., Johnson, K., Ehrlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science*, 12, 83–87. doi:10.1111/1467-8721.01235
- Dunning, D., Leuenberger, A., & Sherman, D. A. (1995). A new look at motivated inference: Are self-serving theories of success a product of motivational forces. *Journal of Personality and Social Psychology*, 69, 58–68. Retrieved from <http://dx.doi.org/10.1037/0022-3514.69.1.58>
- Dunning, D., & McElwee, R. O. (1995). Idiosyncratic trait definitions: Implications for self-description and social judgments. *Journal of Personality and Social Psychology*, 68, 936–946. Retrieved from <http://dx.doi.org/10.1037/0022-3514.68.5.936>
- Dunning, D., Meyerowitz, J. A., & Holzberg, A. D. (1989). Ambiguity and self-evaluation: The role of idiosyncratic trait definitions in self-serving assessments of ability. In T. Gilovich, D. Griffin, & D. Kahneman (Eds.), *Heuristics and biases: The psychology of intuitive judgment* (pp. 324–333). New York, NY: Cambridge University Press. Retrieved from <http://dx.doi.org/10.1017/cbo9780511808098.020>
- Duval, T. S., & Silvia, P. J. (2002). Self-awareness, probability of improvement, and the self-serving bias. *Journal of Personality and Social Psychology*, 82, 49–61. Retrieved from <http://dx.doi.org/10.1037/0022-3514.82.1.49>
- Elliston, J. (1995). *Investigative interviewing training: Report on the effectiveness of skill and procedure from training to the operational environment*. Bedford: Marshgate Consulting Limited.
- Epley, N., & Dunning, D. (2000). Feeling “Holier than thou”: Are self-serving assessments produced by errors in self or social prediction? *Journal of Personality and Social Psychology*, 79, 861–875. Retrieved from <http://dx.doi.org/10.1037/0022-3514.79.6.861>
- Goodyear, R. K. (1997). Psychological expertise and the role of individual differences: An exploration of issues. *Educational Psychology Review*, 9, 251–265.
- Gosling, S. D., John, O., Craik, K. H., & Robins, R. W. (1998). Do people know how they behave? Self-reported act frequencies compared with on-line codings by observers. *Journal of Personality and Social Psychology*, 74, 1337–1349. Retrieved from <http://dx.doi.org/10.1037/0022-3514.74.5.1337>
- Griffiths, A. (2008). An examination into the efficacy of police advanced investigative interview training? (Unpublished PhD thesis). University of Portsmouth, Portsmouth.
- Griffiths, A., & Milne, R. (2006). Will it all end in tiers? Police interviews with suspects in Britain. In T. Williamson (Ed.), *Investigative interviewing: Rights, research and regulation* (pp. 167–189). Cullompton: Willan.
- Gruder, C. L. (1971). Determinants of social comparison choices. *Journal of Experimental Social Psychology*, 7, 473–89. Retrieved from [http://dx.doi.org/10.1016/0022-1031\(71\)90010-2](http://dx.doi.org/10.1016/0022-1031(71)90010-2)
- Hanrahan, S. J., & Isaacs, G. (2001). Assessing self- and peer-assessment: The students’ views. *Higher Education Research and Development*, 20, 53–70. Retrieved from <http://dx.doi.org/10.1080/07294360123776>
- Hayes, A. F., & Dunning, D. (1997). Construal processes and trait ambiguity: Implications for self-peer agreement in personality judgment. *Journal of Personality and Social Psychology*, 72, 644–677. Retrieved from <http://dx.doi.org/10.1037/0022-3514.72.3.664>
- Kahneman, D., Slovic, P., & Tversky, A. (1982). *Judgments under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press. Retrieved from <http://dx.doi.org/10.1017/cbo9780511809477.002>
- Kihlström, J. F., Beer, J. S., & Klein, S. B. (2003). Self and identity as memory. In M. R. Leary & J. P. Tangney (Eds.), *Handbook of self and identity* (pp. 68–90). New York, NY: Guilford Press.
- Klenowski, V. (1995). Student self-evaluation processes in student-centred teaching and learning contexts of Australia and England. *Assessment in Education: Principles, Policy and Practice*, 2, 145–154. Retrieved from <http://dx.doi.org/10.1080/0969594950020203>
- Kruger, J., & Dunning, D. (1999). Unskilled and unaware of it: How difficulties in recognizing one’s own incompetence lead to inflated self-assessments. *Journal of Personality and Social Psychology*, 77, 1121–1134. doi:10.1037/0022-3514.77.6.1121
- La Rooy, D., Lamb, M. E., & Memon, A. (2011). Forensic interviews with children in Scotland: A survey of interview practice among police. *Journal of Police and Criminal Psychology*, 26, 26–34. Retrieved from <http://dx.doi.org/10.1007/s11896-010-9072-9>

- Lamb, M. E., Sternberg, K. J., Orbach, Y., Esplin, P. W., & Mitchell, S. (2002). Is ongoing feedback necessary to maintain the quality of investigative interviews with allegedly abused children? *Applied Developmental Science*, 6, 35–41. Retrieved from http://dx.doi.org/10.1207/s1532480xads0601_04
- Lamb, M. E., Sternberg, K. J., Orbach, Y., Hershkowitz, I., Horowitz, D., & Esplin, P. W. (2000). The effects of intensive training and ongoing supervision on the quality of investigative interviews with alleged sex abuse victims. *Applied Developmental Science*, 6, 114–125. Retrieved from http://dx.doi.org/10.1207/s1532480xads0603_2
- Lockwood, P., & Kunda, Z. (1997). Superstars and me: Predicting the impact of role models on self. *Journal of Personality and Social Psychology*, 73, 91–103. Retrieved from <http://dx.doi.org/10.1037/0022-3514.73.1.91>
- Mamede, S., Schmidt, H. G., & Penaforte, J. C. (2008). Effects of reflective practice on the accuracy of medical diagnoses. *Medical Education*, 42, 468–75. Retrieved from <http://dx.doi.org/10.1111/j.1365-2923.2008.03030.x>
- Marshall, M. A., & Brown, J. D. (2007). On the psychological benefits of self-enhancement. In E. C. Chang (Ed.), *Self-criticism and self-enhancement: Theory, research, and clinical implications* (pp. 19–35). Washington, DC: American Psychological Association. Retrieved from <http://dx.doi.org/10.1037/11624-002>
- Mezulis, A. H., Abramson, L. Y., Hyde, J. S., & Hankin, B. L. (2004). Is there a universal positivity bias in attributions? A meta-analytic review of individual, developmental, and cultural differences in the self-serving attributional bias. *Psychological Bulletin*, 130, 711–747.
- Miller, D. T., Turnbull, W., & McFarland, C. (1988). Particularistic and universalistic evaluation in the social comparison process. *Journal of Personality and Social Psychology*, 55, 908–17. Retrieved from <http://dx.doi.org/10.1037/0022-3514.55.6.908>
- Moon, J. (2006). *Reflection in learning and professional development: Theory & practice*. London: Routledge Falmer. Retrieved from <http://dx.doi.org/10.4324/9780203822296>
- Oettingen, G., & Gollwitzer, P. M. (2001). Goal setting and goal striving. In A. Tesser & N. Schwarz (Eds.), *Intra-individual processes. Volume 1 of the Blackwell handbook in social psychology* (pp. 329–347). Editors-in-chief: M. Hewstone and M. Brewer. Oxford: Blackwell. Retrieved from <http://dx.doi.org/10.1002/9780470998519.ch15>
- O'Rourke, R. (1998). The learning journal: From chaos to coherence. *Assessment and Evaluation in Higher Education*, 23, 403–413. Retrieved from <http://dx.doi.org/10.1080/0260293980230407>
- Paisley, E. (1998). *An investigation into the effectiveness of the Bedfordshire Police Investigative interview Tape Monitoring Procedure* (Unpublished MA dissertation). De Montfort University, Bedford.
- Patton, J. G., Wood, S. J., & Agarenzo, T. (1997). Enhancing the clinical practicum experience through journal writing. *Journal of Nursing Education*, 36, 238–240.
- Powell, M. (2002). Specialist training in investigative and evidential interviewing: Is it having any effect on the behaviour of professionals in the field? *Psychiatry, Psychology and Law*, 9, 44–55. Retrieved from <http://dx.doi.org/10.1375/pplt.2002.9.1.44>
- Powell, M. B., Wright, R., & Hughes-Scholes, C. (2011). Contrasting the perceptions of child testimony experts, prosecutors and police officers regarding individual child abuse interviews. *Psychiatry, Psychology & Law*, 18, 33–43. Retrieved from <http://dx.doi.org/10.1080/13218710903566995>
- Prentice, D. A. (1990). Familiarity and differences in self- and other-representation. *Journal of Personality and Social Psychology*, 59, 369–383. Retrieved from <http://dx.doi.org/10.1037/0022-3514.59.3.369>
- Preuss, G., & Alicke, M. D. (2009). Everybody loves me: Self-evaluations and metaperceptions of dating popularity. *Personality and Social Psychology Bulletin*, 35, 937–950. Retrieved from <http://dx.doi.org/10.1177/0146167209335298>
- Ridley, A., Van Rheede, J., & Wilcock, R. (2015). Interviews, intermediaries and interventions: Barristers', police officers' and mock-jurors' perceptions of a child witness interview. *Investigative Interviewing: Research and Practice*, 7, 21–35.
- Rigg, K. (1999). An Assessment of the positive as compared to the normative in the use of best practice, as defined by the PEACE model of investigative interviewing, in the investigation and

- interviewing of suspects by officers who have been trained in the use of the PEACE model by Northumbria Police (Unpublished MSc dissertation). University of Portsmouth, Portsmouth.
- Roese, N. J., & Olson, J. M. (2007). Better, stronger, faster: Self-serving judgment, affect regulation, and the optimal vigilance hypothesis. *Perspectives on Psychological Science*, 2, 124–141. Retrieved from <http://dx.doi.org/10.1111/j.1745-6916.2007.00033.x>
- Ruth-Sahd, L. A. (2003). Reflective practice: A critical analysis of data-based studies and implications for nursing education. *Journal of Nursing Education*, 42, 488–497.
- Salmon, R. (2006). Assessing reflective learning: precepts, percepts and practices. *Investigations in University Teaching and Learning*, 4, 97–103.
- Schon, D. (1987). *Educating the reflective practitioner*. San Francisco, CA: Jossey Bass. Retrieved from <http://dx.doi.org/10.1002/chp.4750090207>
- Sedikides, C. (1999). A multiplicity of motives: The case of self-improvement. *Psychological Inquiry*, 9, 64–65. Retrieved from http://dx.doi.org/10.1207/s15327965pli1001_10
- Sedikides, C., Campbell, W. K., Reeder, G., & Elliot, A. J. (2002). The self in relationships: Whether, how, and when close others put the self “in its place.” In W. Stroebe & M. Hewstone (Eds.), *European review of social psychology* (Vol. 12, pp. 237–265). Chichester: Wiley. Retrieved from <http://dx.doi.org/10.1080/14792772143000076>
- Sedikides, C., Gaertner, L., & Toguchi, Y. (2003). Pancultural self-enhancement. *Journal of Personality and Social Psychology*, 84, 60–70. Retrieved from <http://dx.doi.org/10.1037/0022-3514.84.1.60>
- Sedikides, C., & Green, J. D. (2006). The mnemonic neglect model: Experimental demonstrations of inhibitory repression in normal adults. *Behavioral and Brain Sciences*, 29, 532–533. Retrieved from <http://dx.doi.org/10.1017/s0140525x06449116>
- Sedikides, C., Green, J. D., & Pinter, B. (2004). Self-protective memory. In D. R. Beike, J. M. Lampinen, & D. A. Behrend (Eds.), *The self and memory* (pp. 161–179). Philadelphia, PA: Psychology Press.
- Sedikides, C., & Gregg, A. P. (2008). Self-enhancement: Food for thought. *Perspectives on Psychological Science*, 3, 102–116. Retrieved from <http://dx.doi.org/10.1111/j.1745-6916.2008.00068.x>
- Sedikides, C., Gregg, A. P., & Hart, C. M. (2008). The importance of being modest. In C. Sedikides, & S. J. Spencer (Eds.), *The self* (pp. 163–184). New York, NY: Psychology Press.
- Sedikides, C., & Gregg, A. P. (2003). Portraits of the self. In M. A. Hogg & J. Cooper (Eds.), *Sage handbook of social psychology* (pp. 110–138). London: Sage Publications. Retrieved from <http://dx.doi.org/10.4135/9781848608221.n5>
- Sedikides, C., & Gregg, A. P. (2006). The self as a point of contact between social psychology and motivation. In P. A. M. Van Lange (Ed.), *Bridging social psychology* (pp. 233–238). Philadelphia, PA: Psychology Press.
- Sedikides, C., & Luke, M. (2007). On when self-enhancement and self-criticism function adaptively and maladaptively. In E. C. Chang (Ed.), *Self-criticism and self-enhancement: Theory, research, and clinical implications* (pp. 181–198). Washington, DC: APA Books. Retrieved from <http://dx.doi.org/10.1037/11624-011>
- Sedikides, C., & Strube, M. J. (1997). Self-evaluation: To thine own self be good, to thine own self be sure, to thine own self be true, and to thine own self be better. In M. P. Zanna (Ed.), *Advances in experimental social psychology* (Vol. 29, pp. 209–269). New York, NY: Academic Press. Retrieved from [http://dx.doi.org/10.1016/s0065-2601\(08\)60018-0](http://dx.doi.org/10.1016/s0065-2601(08)60018-0)
- Shawyer, A. (2009) Investigative interviewing: Interviewing counter fraud and deception (Unpublished PhD thesis). University of Portsmouth, Portsmouth.
- Sheikh, A. I., Milne, D. L., & MacGregor, B. V. (2007). A model of personal professional development in the systematic training of clinical psychologists. *Clinical Psychology and Psychotherapy*, 14, 278–287. Retrieved from <http://dx.doi.org/10.1002/cpp.540>
- Shepherd, E., & Griffiths, A. (2013). *Investigative interviewing*. Oxford: Oxford University Press.
- Skowronski, J. J., Betz, A. L., Thompson, C. P., & Shannon, L. (1991). Social memory in everyday life: The recall of self-events and other-events. *Journal of Personality and Social Psychology*, 60, 831–843. Retrieved from <http://dx.doi.org/10.1037/0022-3514.60.6.831>
- Skowronski, J. J., Gibbons, J. A., Vogl, R. J., & Walker, W. R. (2004). The effect of social disclosure on the affective intensity provoked by autobiographical memories. *Self and Identity*, 3, 285–309. Retrieved from <http://dx.doi.org/10.1080/13576500444000065>

- Stefani, L. A. J. (1998). Assessment in partnership with learners. *Assessment and Evaluation in Higher Education*, 23, 339–350. Retrieved from <http://dx.doi.org/10.1080/0260293980230402>
- Stevens, A. (1998). An evaluation of the process of developing interview skills for uniform constables within Humberside police (Unpublished MEd dissertation). University of Hull, Hull.
- Stockdale, J. (1993). *Management and supervision of police interviews*. Police Research Group Paper 5. London: Home Office.
- Suls, J., & Wills, T. A. (Eds.). (1991) *Social comparison: Contemporary theory and research*. Hillsdale, NJ: Erlbaum.
- Taylor, E. (2014). Honour among thieves? How morality and rationality influence the decision-making processes of convicted domestic burglars. *Criminology and Criminal Justice: An International Journal*, 14, 487–502. Retrieved from <http://dx.doi.org/10.1177/1748895813505232>
- Taylor, S. E., Lerner, J. S., Sherman, D. K., Sage, R. M., & McDowell, N. K. (2003). Portrait of the self-enhancer: Well-adjusted and well-liked or maladjusted and friendless? *Journal of Personality and Social Psychology*, 84, 165–176. Retrieved from <http://dx.doi.org/10.1037/0022-3514.84.1.165>
- Van Overwalle, F., & Labiouse, C. (2004). A recurrent connectionist model of person impression formation. *Personality and Social Psychology Review*, 8, 28–61. Retrieved from http://dx.doi.org/10.1207/s15327957pspr0801_2
- Walker, W. R., Skowronski, J. J., & Thompson, C. P. (2003). Life is pleasant – and memory helps to keep it that way. *Review of General Psychology*, 7, 203–210. Retrieved from <http://dx.doi.org/10.1037/1089-2680.7.2.203>
- Walsh, D., & Bull, R. (2010). What really is effective in interviews with suspects? A study comparing interviewing skills against interviewing outcomes. *Legal and Criminological Psychology*, 15, 305–321. Retrieved from <http://dx.doi.org/10.1348/135532509x463356>
- Walsh, D., & Bull, R. (2011). Benefit fraud investigative interviewing: A self-report study of investigation professionals' beliefs concerning practice. *Journal of Investigative Psychology and Offender Profiling*, 8, 131–148. Retrieved from <http://dx.doi.org/10.1002/jip.137>
- Walsh, D., & Bull, R. (2012). Examining rapport in investigative interviews with suspects: Does its building and maintenance work?. *Journal of Police and Criminal Psychology*, 27, 73–84.
- Walsh, D., & Bull, R. (2015). The association between evidence disclosure strategies, questioning strategies, and interview outcomes. *Psychology, Crime and Law*, 21, 661–680. Retrieved from <http://dx.doi.org/10.1080/1068316x.2015.1028544>
- Walsh, D., & Milne, R. (2007). Perceptions of benefit fraud staff in the UK: Giving PEACE a chance?. *Public Administration*, 85, 525–540. doi:10.1111/j.1467-9299.2007.00645.x
- Walsh, D., & Milne, R. (2008). Keeping the PEACE? A study of investigative interviewing practices in the public sector. *Legal and Criminological Psychology*, 13, 39–57. Retrieved from <http://dx.doi.org/10.1348/135532506x157179>
- Walsh, D., Oxburgh, G., Redlich, A., & Myklebust, T. (Eds.). (2016). *International developments and practices in investigative interviewing and interrogation (Volume 2)*. London: Routledge.
- Weiner, B. (1972). *Theories of motivation: From mechanism to cognition*. Chicago, IL: Rand.
- Weinstein, N. D. (1984). Why it won't happen to me: Perceptions of risk factors and susceptibility. *Health Psychology*, 3, 431–457. doi:10.1037/0278-6133.3.5.431. .
- Wheeler, L. (1996). Motivation as a determinant of upward social comparison. *Journal of Experimental Social Psychology*, 2, 27–31. doi:10.1016/0022-1031(66)90062-x
- Wilson, T. D., & Dunn, E. W. (2004). Self-knowledge: Its limits, value, and potential for improvement. *Annual Review of Psychology*, 55, 493–518. doi:10.1146/annurev.psych.55.090902.141954
- Wood, J. V. (1989). Theory and research concerning social comparison of personality attributes. *Psychological Bulletin*, 106, 231–48. doi:10.1037/0033-2909.106.2.231
- Woodward, H. (1998). Reflective journals and portfolios: Learning through assessment. *Assessment and Evaluation in Higher Education*, 23, 415–423. Retrieved from <http://dx.doi.org/10.1080/0260293980230408>